

- 1 -

SEQUENCE LISTING

<110> F. Hoffmann-La Roche AG

<120> Method for the recombinant expression of an N-terminal fragment of hepatocyte growth factor

<130> 22388 WO

<150> EP 04004951.2

<151> 2004-03-03

<160> 4

<170> PatentIn version 3.2

<210> 1

<211> 1389

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(1389)

<223> DNA sequence encoding the alpha-chain of hepatocyte growth factor (HGF)

<400> 1

caa agg aaa aga aga aat aca att cat gaa ttc aaa aaa tca gca aag	48
Gln Arg Lys Arg Arg Asn Thr Ile His Glu Phe Lys Lys Ser Ala Lys	
1 5 10 15	
act acc cta atc aaa ata gat cca gca ctg aag ata aaa acc aaa aaa	96
Thr Thr Leu Ile Lys Ile Asp Pro Ala Leu Lys Ile Lys Thr Lys Lys	
20 25 30	
gtg aat act gca gac caa tgt gct aat aga tgt act agg aat aaa gga	144
Val Asn Thr Ala Asp Gln Cys Ala Asn Arg Cys Thr Arg Asn Lys Gly	
35 40 45	
ctt cca ttc act tgc aag gct ttt gtt ttt gat aaa gca aga aaa caa	192
Leu Pro Phe Thr Cys Lys Ala Phe Val Phe Asp Lys Ala Arg Lys Gln	
50 55 60	
tgc ctc tgg ttc ccc ttc aat agc atg tca agt gga gtg aaa aaa gaa	240
Cys Leu Trp Phe Pro Phe Asn Ser Met Ser Ser Gly Val Lys Lys Glu	
65 70 75 80	
ttt ggc cat gaa ttt gac ctc tat gaa aac aaa gac tac att aga aac	288
Phe Gly His Glu Phe Asp Leu Tyr Glu Asn Lys Asp Tyr Ile Arg Asn	
85 90 95	

- 2 -

tgc atc att ggt aaa gga cgc agc tac aag gga aca gta tct atc act	336
Cys Ile Ile Gly Lys Gly Arg Ser Tyr Lys Gly Thr Val Ser Ile Thr	
100 105 110	
aag agt ggc atc aaa tgt cag ccc tgg agt tcc atg ata cca cac gaa	384
Lys Ser Gly Ile Lys Cys Gln Pro Trp Ser Ser Met Ile Pro His Glu	
115 120 125	
cac agc ttt ttg cct tcg agc tat cgg ggt aaa gac cta cag gaa aac	432
His Ser Phe Leu Pro Ser Ser Tyr Arg Gly Lys Asp Leu Gln Glu Asn	
130 135 140	
tac tgt cga aat cct cga ggg gaa gaa ggg gga ccc tgg tgt ttc aca	480
Tyr Cys Arg Asn Pro Arg Gly Glu Glu Gly Gly Pro Trp Cys Phe Thr	
145 150 155 160	
agc aat cca gag gta cgc tac gaa gtc tgt gac att cct cag tgt tca	528
Ser Asn Pro Glu Val Arg Tyr Glu Val Cys Asp Ile Pro Gln Cys Ser	
165 170 175	
gaa gtt gaa tgc atg acc tgc aat ggg gag agt tat cga ggt ctc atg	576
Glu Val Glu Cys Met Thr Cys Asn Gly Glu Ser Tyr Arg Gly Leu Met	
180 185 190	
gat cat aca gaa tca ggc aag att tgt cag cgc tgg gat cat cag aca	624
Asp His Thr Glu Ser Gly Lys Ile Cys Gln Arg Trp Asp His Gln Thr	
195 200 205	
cca cac cgg cac aaa ttc ttg cct gaa aga tat ccc gac aag ggc ttt	672
Pro His Arg His Lys Phe Leu Pro Glu Arg Tyr Pro Asp Lys Gly Phe	
210 215 220	
gat gat aat tat tgc cgc aat ccc gat ggc cag ccg agg cca tgg tgc	720
Asp Asp Asn Tyr Cys Arg Asn Pro Asp Gly Gln Pro Arg Pro Trp Cys	
225 230 235 240	
tat act ctt gac cct cac acc cgc tgg gag tac tgt gca att aaa aca	768
Tyr Thr Leu Asp Pro His Thr Arg Trp Glu Tyr Cys Ala Ile Lys Thr	
245 250 255	
tgc gct gac aat act atg aat gac act gat gtt cct ttg gaa aca act	816
Cys Ala Asp Asn Thr Met Asn Asp Thr Asp Val Pro Leu Glu Thr Thr	
260 265 270	
gaa tgc atc caa ggt caa gga gaa ggc tac agg ggc act gtc aat acc	864
Glu Cys Ile Gln Gly Gln Gly Glu Gly Tyr Arg Gly Thr Val Asn Thr	
275 280 285	
att tgg aat gga att cca tgt cag cgt tgg gat tct cag tat cct cac	912
Ile Trp Asn Gly Ile Pro Cys Gln Arg Trp Asp Ser Gln Tyr Pro His	
290 295 300	
gag cat gac atg act cct gaa aat ttc aag tgc aag gac cta cga gaa	960
Glu His Asp Met Thr Pro Glu Asn Phe Lys Cys Lys Asp Leu Arg Glu	
305 310 315 320	

- 3 -

aat tac tgc cga aat cca gat ggg tct gaa tca ccc tgg tgt ttt acc	1008
Asn Tyr Cys Arg Asn Pro Asp Gly Ser Glu Ser Pro Trp Cys Phe Thr	
325 330 335	
act gat cca aac atc cga gtt ggc tac tgc tcc caa att cca aac tgt	1056
Thr Asp Pro Asn Ile Arg Val Gly Tyr Cys Ser Gln Ile Pro Asn Cys	
340 345 350	
gat atg tca cat gga caa gat tgt tat cgt ggg aat ggc aaa aat tat	1104
Asp Met Ser His Gly Gln Asp Cys Tyr Arg Gly Asn Gly Lys Asn Tyr	
355 360 365	
atg ggc aac tta tcc caa aca aga tct gga cta aca tgt tca atg tgg	1152
Met Gly Asn Leu Ser Gln Thr Arg Ser Gly Leu Thr Cys Ser Met Trp	
370 375 380	
gac aag aac atg gaa gac tta cat cgt cat atc ttc tgg gaa cca gat	1200
Asp Lys Asn Met Glu Asp Leu His Arg His Ile Phe Trp Glu Pro Asp	
385 390 395 400	
gca agt aag ctg aat gag aat tac tgc cga aat cca gat gat gat gct	1248
Ala Ser Lys Leu Asn Glu Asn Tyr Cys Arg Asn Pro Asp Asp Asp Ala	
405 410 415	
cat gga ccc tgg tgc tac acg gga aat cca ctc att cct tgg gat tat	1296
His Gly Pro Trp Cys Tyr Thr Gly Asn Pro Leu Ile Pro Trp Asp Tyr	
420 425 430	
tgc cct att tct cgt tgt gaa ggt gat acc aca cct aca ata gtc aat	1344
Cys Pro Ile Ser Arg Cys Glu Gly Asp Thr Thr Pro Thr Ile Val Asn	
435 440 445	
tta gac cat ccc gta ata tct tgt gcc aaa acg aaa caa ttg cga	1389
Leu Asp His Pro Val Ile Ser Cys Ala Lys Thr Lys Gln Leu Arg	
450 455 460	
<210> 2	
<211> 463	
<212> PRT	
<213> Homo sapiens	
<400> 2	
Gln Arg Lys Arg Arg Asn Thr Ile His Glu Phe Lys Lys Ser Ala Lys	
1 5 10 15	
Thr Thr Leu Ile Lys Ile Asp Pro Ala Leu Lys Ile Lys Thr Lys Lys	
20 25 30	
Val Asn Thr Ala Asp Gln Cys Ala Asn Arg Cys Thr Arg Asn Lys Gly	
35 40 45	
Leu Pro Phe Thr Cys Lys Ala Phe Val Phe Asp Lys Ala Arg Lys Gln	
50 55 60	

- 4 -

Cys Leu Trp Phe Pro Phe Asn Ser Met Ser Ser Gly Val Lys Lys Glu
 65 70 75 80

Phe Gly His Glu Phe Asp Leu Tyr Glu Asn Lys Asp Tyr Ile Arg Asn
 85 90 95

Cys Ile Ile Gly Lys Gly Arg Ser Tyr Lys Gly Thr Val Ser Ile Thr
 100 105 110

Lys Ser Gly Ile Lys Cys Gln Pro Trp Ser Ser Met Ile Pro His Glu
 115 120 125

His Ser Phe Leu Pro Ser Ser Tyr Arg Gly Lys Asp Leu Gln Glu Asn
 130 135 140

Tyr Cys Arg Asn Pro Arg Gly Glu Glu Gly Gly Pro Trp Cys Phe Thr
 145 150 155 160

Ser Asn Pro Glu Val Arg Tyr Glu Val Cys Asp Ile Pro Gln Cys Ser
 165 170 175

Glu Val Glu Cys Met Thr Cys Asn Gly Glu Ser Tyr Arg Gly Leu Met
 180 185 190

Asp His Thr Glu Ser Gly Lys Ile Cys Gln Arg Trp Asp His Gln Thr
 195 200 205

Pro His Arg His Lys Phe Leu Pro Glu Arg Tyr Pro Asp Lys Gly Phe
 210 215 220

Asp Asp Asn Tyr Cys Arg Asn Pro Asp Gly Gln Pro Arg Pro Trp Cys
 225 230 235 240

Tyr Thr Leu Asp Pro His Thr Arg Trp Glu Tyr Cys Ala Ile Lys Thr
 245 250 255

Cys Ala Asp Asn Thr Met Asn Asp Thr Asp Val Pro Leu Glu Thr Thr
 260 265 270

Glu Cys Ile Gln Gly Gln Gly Glu Gly Tyr Arg Gly Thr Val Asn Thr
 275 280 285

- 5 -

Ile Trp Asn Gly Ile Pro Cys Gln Arg Trp Asp Ser Gln Tyr Pro His
 290 295 300

Glu His Asp Met Thr Pro Glu Asn Phe Lys Cys Lys Asp Leu Arg Glu
 305 310 315 320

Asn Tyr Cys Arg Asn Pro Asp Gly Ser Glu Ser Pro Trp Cys Phe Thr
 325 330 335

Thr Asp Pro Asn Ile Arg Val Gly Tyr Cys Ser Gln Ile Pro Asn Cys
 340 345 350

Asp Met Ser His Gly Gln Asp Cys Tyr Arg Gly Asn Gly Lys Asn Tyr
 355 360 365

Met Gly Asn Leu Ser Gln Thr Arg Ser Gly Leu Thr Cys Ser Met Trp
 370 375 380

Asp Lys Asn Met Glu Asp Leu His Arg His Ile Phe Trp Glu Pro Asp
 385 390 395 400

Ala Ser Lys Leu Asn Glu Asn Tyr Cys Arg Asn Pro Asp Asp Ala
 405 410 415

His Gly Pro Trp Cys Tyr Thr Gly Asn Pro Leu Ile Pro Trp Asp Tyr
 420 425 430

Cys Pro Ile Ser Arg Cys Glu Gly Asp Thr Thr Pro Thr Ile Val Asn
 435 440 445

Leu Asp His Pro Val Ile Ser Cys Ala Lys Thr Lys Gln Leu Arg
 450 455 460

<210> 3
 <211> 1350
 <212> DNA
 <213> Artificial

<220>
 <223> dna coding for NK4

<220>
 <221> CDS
 <222> (1)..(1350)

<400> 3

- 6 -

atg tct cgt aaa cgt cgt aat act att cat gaa ttc aaa aaa tca gca	48
Met Ser Arg Lys Arg Arg Asn Thr Ile His Glu Phe Lys Lys Ser Ala	
1 5 10 15	
aag act acc cta atc aaa ata gat cca gca ctg aag ata aaa acc aaa	96
Lys Thr Thr Leu Ile Lys Ile Asp Pro Ala Leu Lys Ile Lys Thr Lys	
20 25 30	
aaa gtg aat act gca gac caa tgt gct aat aga tgt act agg aat aaa	144
Lys Val Asn Thr Ala Asp Gln Cys Ala Asn Arg Cys Thr Arg Asn Lys	
35 40 45	
gga ctt cca ttc act tgc aag gct ttt gtt ttt gat aaa gca aga aaa	192
Gly Leu Pro Phe Thr Cys Lys Ala Phe Val Phe Asp Lys Ala Arg Lys	
50 55 60	
caa tgc ctc tgg ttc ccc ttc aat agc atg tca agt gga gtg aaa aaa	240
Gln Cys Leu Trp Phe Pro Phe Asn Ser Met Ser Ser Gly Val Lys Lys	
65 70 75 80	
gaa ttt ggc cat gaa ttt gac ctc tat gaa aac aaa gac tac att aga	288
Glu Phe Gly His Glu Phe Asp Leu Tyr Glu Asn Lys Asp Tyr Ile Arg	
85 90 95	
aac tgc atc att ggt aaa gga cgc agc tac aag gga aca gta tct atc	336
Asn Cys Ile Ile Gly Lys Gly Arg Ser Tyr Lys Gly Thr Val Ser Ile	
100 105 110	
act aag agt ggc atc aaa tgt cag ccc tgg agt tcc atg ata cca cac	384
Thr Lys Ser Gly Ile Lys Cys Gln Pro Trp Ser Ser Met Ile Pro His	
115 120 125	
gaa cac agc ttt ttg cct tcg agc tat cgg ggt aaa gac cta cag gaa	432
Glu His Ser Phe Leu Pro Ser Ser Tyr Arg Gly Lys Asp Leu Gln Glu	
130 135 140	
aac tac tgt cga aat cct cga ggg gaa gaa ggg gga ccc tgg tgt ttc	480
Asn Tyr Cys Arg Asn Pro Arg Gly Glu Glu Gly Gly Pro Trp Cys Phe	
145 150 155 160	
aca agc aat cca gag gta cgc tac gaa gtc tgt gac att cct cag tgt	528
Thr Ser Asn Pro Glu Val Arg Tyr Glu Val Cys Asp Ile Pro Gln Cys	
165 170 175	
tca gaa gtt gaa tgc atg acc tgc aat ggg gag agt tat cga ggt ctc	576
Ser Glu Val Glu Cys Met Thr Cys Asn Gly Glu Ser Tyr Arg Gly Leu	
180 185 190	
atg gat cat aca gaa tca ggc aag att tgt cag cgc tgg gat cat cag	624
Met Asp His Thr Glu Ser Gly Lys Ile Cys Gln Arg Trp Asp His Gln	
195 200 205	

- 7 -

aca cca cac cgg cac aaa ttc ttg cct gaa aga tat ccc gac aag ggc Thr Pro His Arg His Lys Phe Leu Pro Glu Arg Tyr Pro Asp Lys Gly 210 215 220	672
ttt gat gat aat tat tgc cgc aat ccc gat ggc cag ccg agg cca tgg Phe Asp Asp Asn Tyr Cys Arg Asn Pro Asp Gly Gln Pro Arg Pro Trp 225 230 235 240	720
tgc tat act ctt gac cct cac acc cgc tgg gag tac tgt gca att aaa Cys Tyr Thr Leu Asp Pro His Thr Arg Trp Glu Tyr Cys Ala Ile Lys 245 250 255	768
aca tgc gct gac aat act atg aat gac act gat gtt cct ttg gaa aca Thr Cys Ala Asp Asn Thr Met Asn Asp Thr Asp Val Pro Leu Glu Thr 260 265 270	816
act gaa tgc atc caa ggt caa gga gaa ggc tac agg ggc act gtc aat Thr Glu Cys Ile Gln Gly Gln Gly Glu Gly Tyr Arg Gly Thr Val Asn 275 280 285	864
acc att tgg aat gga att cca tgt cag cgt tgg gat tct cag tat cct Thr Ile Trp Asn Gly Ile Pro Cys Gln Arg Trp Asp Ser Gln Tyr Pro 290 295 300	912
cac gag cat gac atg act cct gaa aat ttc aag tgc aag gac cta cga His Glu His Asp Met Thr Pro Glu Asn Phe Lys Cys Lys Asp Leu Arg 305 310 315 320	960
gaa aat tac tgc cga aat cca gat ggg tct gaa tca ccc tgg tgt ttt Glu Asn Tyr Cys Arg Asn Pro Asp Gly Ser Glu Ser Pro Trp Cys Phe 325 330 335	1008
acc act gat cca aac atc cga gtt ggc tac tgc tcc caa att cca aac Thr Thr Asp Pro Asn Ile Arg Val Gly Tyr Cys Ser Gln Ile Pro Asn 340 345 350	1056
tgt gat atg tca cat gga caa gat tgt tat cgt ggg aat ggc aaa aat Cys Asp Met Ser His Gly Gln Asp Cys Tyr Arg Gly Asn Gly Lys Asn 355 360 365	1104
tat atg ggc aac tta tcc caa aca aga tct gga cta aca tgt tca atg Tyr Met Gly Asn Leu Ser Gln Thr Arg Ser Gly Leu Thr Cys Ser Met 370 375 380	1152
tgg gac aag aac atg gaa gac tta cat cgt cat atc ttc tgg gaa cca Trp Asp Lys Asn Met Glu Asp Leu His Arg His Ile Phe Trp Glu Pro 385 390 395 400	1200
gat gca agt aag ctg aat gag aat tac tgc cga aat cca gat gat gat Asp Ala Ser Lys Leu Asn Glu Asn Tyr Cys Arg Asn Pro Asp Asp Asp 405 410 415	1248
gct cat gga ccc tgg tgc tac acg gga aat cca ctc att cct tgg gat Ala His Gly Pro Trp Cys Tyr Thr Gly Asn Pro Leu Ile Pro Trp Asp 420 425 430	1296

- 8 -

tat tgc cct att tct cgt tgt gaa ggt gat acc aca cct aca atc gtt	1344
Tyr Cys Pro Ile Ser Arg Cys Glu Gly Asp Thr Thr Pro Thr Ile Val	
435 440 445	

taa tag	1350
---------	------

<210> 4
 <211> 448
 <212> PRT
 <213> Artificial

<220>
 <223> protein sequence of NK4

<400> 4

Met Ser Arg Lys Arg Arg Asn Thr Ile His Glu Phe Lys Lys Ser Ala
1 5 10 15

Lys Thr Thr Leu Ile Lys Ile Asp Pro Ala Leu Lys Ile Lys Thr Lys
20 25 30

Lys Val Asn Thr Ala Asp Gln Cys Ala Asn Arg Cys Thr Arg Asn Lys
35 40 45

Gly Leu Pro Phe Thr Cys Lys Ala Phe Val Phe Asp Lys Ala Arg Lys
50 55 60

Gln Cys Leu Trp Phe Pro Phe Asn Ser Met Ser Ser Gly Val Lys Lys
65 70 75 80

Glu Phe Gly His Glu Phe Asp Leu Tyr Glu Asn Lys Asp Tyr Ile Arg
85 90 95

Asn Cys Ile Ile Gly Lys Gly Arg Ser Tyr Lys Gly Thr Val Ser Ile
100 105 110

Thr Lys Ser Gly Ile Lys Cys Gln Pro Trp Ser Ser Met Ile Pro His
115 120 125

Glu His Ser Phe Leu Pro Ser Ser Tyr Arg Gly Lys Asp Leu Gln Glu
130 135 140

- 9 -

Asn	Tyr	Cys	Arg	Asn	Pro	Arg	Gly	Glu	Glu	Gly	Gly	Pro	Trp	Cys	Phe
145					150					155					160
Thr	Ser	Asn	Pro	Glu	Val	Arg	Tyr	Glu	Val	Cys	Asp	Ile	Pro	Gln	Cys
				165					170					175	
Ser	Glu	Val	Glu	Cys	Met	Thr	Cys	Asn	Gly	Glu	Ser	Tyr	Arg	Gly	Leu
			180					185					190		
Met	Asp	His	Thr	Glu	Ser	Gly	Lys	Ile	Cys	Gln	Arg	Trp	Asp	His	Gln
		195					200					205			
Thr	Pro	His	Arg	His	Lys	Phe	Leu	Pro	Glu	Arg	Tyr	Pro	Asp	Lys	Gly
	210					215					220				
Phe	Asp	Asp	Asn	Tyr	Cys	Arg	Asn	Pro	Asp	Gly	Gln	Pro	Arg	Pro	Trp
225					230					235					240
Cys	Tyr	Thr	Leu	Asp	Pro	His	Thr	Arg	Trp	Glu	Tyr	Cys	Ala	Ile	Lys
				245					250					255	
Thr	Cys	Ala	Asp	Asn	Thr	Met	Asn	Asp	Thr	Asp	Val	Pro	Leu	Glu	Thr
			260					265					270		
Thr	Glu	Cys	Ile	Gln	Gly	Gln	Gly	Glu	Gly	Tyr	Arg	Gly	Thr	Val	Asn
		275					280					285			
Thr	Ile	Trp	Asn	Gly	Ile	Pro	Cys	Gln	Arg	Trp	Asp	Ser	Gln	Tyr	Pro
	290					295					300				
His	Glu	His	Asp	Met	Thr	Pro	Glu	Asn	Phe	Lys	Cys	Lys	Asp	Leu	Arg
305					310					315					320
Glu	Asn	Tyr	Cys	Arg	Asn	Pro	Asp	Gly	Ser	Glu	Ser	Pro	Trp	Cys	Phe
				325					330					335	
Thr	Thr	Asp	Pro	Asn	Ile	Arg	Val	Gly	Tyr	Cys	Ser	Gln	Ile	Pro	Asn
			340					345					350		
Cys	Asp	Met	Ser	His	Gly	Gln	Asp	Cys	Tyr	Arg	Gly	Asn	Gly	Lys	Asn
		355					360					365			
Tyr	Met	Gly	Asn	Leu	Ser	Gln	Thr	Arg	Ser	Gly	Leu	Thr	Cys	Ser	Met
	370					375					380				

- 10 -

Trp Asp Lys Asn Met Glu Asp Leu His Arg His Ile Phe Trp Glu Pro
385 390 395 400

Asp Ala Ser Lys Leu Asn Glu Asn Tyr Cys Arg Asn Pro Asp Asp Asp
405 410 415

Ala His Gly Pro Trp Cys Tyr Thr Gly Asn Pro Leu Ile Pro Trp Asp
420 425 430

Tyr Cys Pro Ile Ser Arg Cys Glu Gly Asp Thr Thr Pro Thr Ile Val
435 440 445